Defence Research and Development Canada
Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

A signal detection model of compound decision tasks

Matthew Duncan

Defence R&D Canada
Technical Report
DRDC Toronto TR 2006-256
December 2006

Canada

A signal detection model of compound decision tasks

Matthew Duncan

Defence R&D Canada – Toronto

Technical Report

DRDC Toronto TR 2006-256

December 2006

Author

Original approved by

Dr. Matthew Duncan

Approved by

Original approved by

Dr. Joe Baranski

Head, Collaborative Behaviour and Integrated Learning Section

Approved for release by

Original approved by

K. M. Sutton

Chair, Document Review and Library Committee

This research was conducted under Thrust code 16ki

Abstract

Detection and identification represent two fundamental types of decision tasks. Although research has focused on each in isolation, the pure forms of these tasks are generally not representative of more complex naturalistic decision environments. For example, a decision maker involved in a Search and Rescue (SAR) operation is faced with locating and identifying a crash site. This kind of decision environment is characterized by both detection and identification components. That is, the decision maker is confronted with uncertainty regarding the presence of a target crash site, and the task of identifying the target from among similar looking structures in the terrain. Decision research using compound decision tasks (detection plus identification) has the advantage of making greater contact with naturalistic environments, but carries with it the cost of increased complexity in analyzing and understanding the data. Because compound decision tasks have more than one locus where decision making can be affected, a formal method is needed to disambiguate (deconfound) effects on decision making and simplify an understanding of decision making performance in complex tasks. In this report a formal model of compound decision tasks (SDT-CD) is presented which fulfills this role. The model was assessed by an analysis of several demonstration data sets from a wide variety of content domains which highlight its ability to simplify the complexity of the task and provide readily interpretable results. In addition to measures of performance and decision bias, the model can be used to test hypotheses about decision making and permits an assessment of whether decision making is optimal.

Résumé

La détection et l'identification représentent deux types fondamentaux de tâches décisionnelles. Ces deux tâches, prises isolément, ont fait l'objet de travaux de recherche. Cependant, dans leur forme pure, elles ne sont pas représentatives, en général, des environnements décisionnels du monde réel, beaucoup plus complexes. Par exemple, un décideur qui participe à une opération de recherche et de sauvetage (SAR) est chargé de localiser et d'identifier le site de l'écrasement d'un aéronef. Ce type d'environnement décisionnel comporte un élément « détection » et un élément « identification ». Cela signifie que le décideur est confronté à une incertitude en ce qui concerne le site de l'écrasement, et qu'il doit identifier le site parmi d'autres structures d'aspect similaire sur le terrain. La recherche sur les tâches décisionnelles complexes (détection plus identification) a l'avantage d'être plus en phase avec le monde réel, mais en contrepartie, il y a le coût de la complexité accrue de l'analyse et de la compréhension des données. Étant donné que les tâches décisionnelles complexes font intervenir plusieurs facteurs susceptibles d'influencer la prise de décision, il faut une méthode formelle pour distinguer (clarifier) les effets des divers facteurs, et pour simplifier l'évaluation des résultats de la de prise de décision. Le rapport présente un modèle de tâches décisionnelles complexes (SDT-CD) qui joue précisément ce rôle. Ce modèle a été évalué grâce à l'analyse de plusieurs ensembles de données de démonstration provenant d'une grande variété de domaines, et il a fait la preuve qu'il est capable de dénouer des tâches complexes et de fournir des résultats faciles à interpréter. En plus de mesurer la performance et la partialité du processus de prise de décision, le modèle peut être utilisé pour tester des hypothèses et pour évaluer si la prise de décision est optimale.

Executive summary

Background

Much of the basic research and theory development on decision making has focused on simple (or pure) decision problems in strict isolation of one another. However, military decision makers outside the laboratory are faced with more complex decision environments. These more naturalistic settings are often characterized by a combination of simple decision problems that need to be addressed simultaneously by the decision maker. For example, Search and Rescue (SAR) operations require identification of a target crash site (identification decision problem) in the context of uncertainty regarding the presence of the target at any given time (detection decision problem). The combination of identification and detection problems form what is referred to as a compound decision task; one that represents a closer approximation to natural decision environments. This research was conducted to extend decision theory into more complex domains of the compound decision task in an effort to provide researchers with the tools required to study decision making that is more directly relevant to the CF. To meet this requirement, a formal model was developed to provide a means of simplifying the assessment of decision making performance, bias, and optimality in various compound decision environments.

Results

The viability of the model was determined by fitting it to several data sets from compound decision tasks across several domains; including SAR, eyewitness identification, and medical diagnosis of tumours.

- Analytically, the model showed a statistically good fit to three of the four demonstration data sets. Overall, the model showed a good qualitative fit to all four data sets.

- Theoretically, the model was successfully used to test hypotheses about the data regarding decision performance, bias, and optimal use of decision rules.

- The model allowed for a simplification and deconfounding of data analysis found to occur with more conventional analysis of compound decision tasks.

Significance

The existence of this model creates greater opportunity for the use of more complex decision tasks in military decision making research. By simplifying and deconfounding data analysis, as well as providing a clear separation of performance, bias, and optimality, defence research can move more easily toward more naturalistic decision settings. This will further strengthen the ties between decision research and direct military applications. Because the model represents an abstract conceptual framework, it can be applied to many different content domains, in addition to dealing with global decision issues that permeate all content domains (e.g., confidence calibration). This will permit decision research to be conducted within a wider range of areas and applications, including an extension of the model into automated decision systems research.

Duncan, M. J. 2006. A Signal Detection Model of Compound Decision Tasks. DRDC Toronto TR 2006-256 Defence R&D Canada.

Sommaire

Contexte

Jusqu'ici, une grande partie des travaux de recherche et des études théoriques sur la prise de décision ont porté sur des problèmes décisionnels simples, pris isolément. Cependant, à l'extérieur des laboratoires, les décideurs militaires font face à des environnements décisionnels beaucoup plus complexes. Ces environnements plus proches du monde réel se caractérisent bien souvent par une combinaison de problèmes décisionnels simples qui doivent être examinés simultanément par le décideur. Par exemple, les opérations de recherche et de sauvetage (SAR) nécessitent l'identification du site de l'écrasement d'un aéronef (problème d'identification), et ils comportent une incertitude quant à l'emplacement de ce site (problème de détection). La combinaison de ces problèmes d'identification et de détection forme ce que l'on appelle une tâche décisionnelle complexe, plus proche de ce que l'on retrouve dans la réalité. Cette recherche a été menée pour étendre la théorie de la prise de décision aux tâches décisionnelles complexes, afin de fournir aux chercheurs les outils dont ils ont besoin pour étudier les problèmes de prise de décision qui intéressent plus particulièrement les FC. Pour répondre à ce besoin, un modèle a été élaboré qui simplifie l'évaluation de la performance, de la partialité et de l'optimalité du processus de prise de décision dans divers environnements décisionnels complexes.

Résultats

Pour évaluer la viabilité du modèle, on l'a appliqué à plusieurs ensembles de données provenant de tâches décisionnelles complexes dans plusieurs domaines d'activité, y compris la recherche et le sauvetage (SAR), l'identification des témoins et le diagnostic médical des tumeurs.

- Sur le plan analytique, le modèle a affiché de bons résultats statistiques pour trois des quatre ensembles de données de démonstration. Globalement, le modèle a affiché de bons résultats qualitatifs pour les quatre ensembles de données.

- Sur le plan théorique, le modèle a été utilisé avec succès pour tester des hypothèses sur les données concernant la performance, la partialité et l'optimalité du processus de prise de décision.

- Le modèle a permis de simplifier et de clarifier les données concernant les tâches décisionnelles complexes, par rapport aux méthodes d'analyse plus conventionnelles.

Signification

Ce modèle crée de nouvelles possibilités pour l'utilisation de tâches décisionnelles plus complexes dans la recherche sur le processus de prise de décision militaire. En simplifiant et en clarifiant l'analyse des données, et en faisant clairement la distinction entre la performance, la partialité et l'optimalité, la recherche militaire pourra examiner plus facilement la prise de décision dans des situations réelles. Cela renforcera les liens entre la recherche sur la prise de décision et les applications militaires réelles. Étant donné que le modèle représente un cadre conceptuel abstrait, il peut être appliqué à de nombreux domaines d'activité, en plus des problèmes décisionnels généraux communs à tous les domaines (ex. : calibrage de la confiance). Cela nous permettra de faire de la recherche dans une grande variété de domaines, et d'appliquer notamment le modèle à la prise de décision assistée par ordinateur.

Duncan, M. J. 2006. Détection des signaux – Modèle de tâches décisionnelles complexes. TR 2006-256 R & D pour la défense Canada.

Table of contents

List of tables

This page intentionally left blank.

1. Introduction

This report focuses on a fundamental problem in decision research regarding methods for analyzing and interpreting data derived from compound decision tasks. The term compound decision task refers to the case where a human (or machine) operator is attempting to identify a target in situations where the target itself may or may not be present at any given time within the array of information presented to the operator. That is, a case of target identification under conditions of uncertainty about the actual presence of the target. There are numerous examples of military tasks which can be categorized as this kind of compound decision task. In Search and Rescue (SAR) operations, search operators will be engaged in visual scanning of an area thought to contain a crash site. Doctors and technicians frequently examine x-rays looking for anomalies. Commanders in a Network Centric Warfare (NCW) environment may have to comb through copious streams of incoming data looking for potential threats or targets. A United Nations Military Observer (UNMO) may be required to help identify a suspected human rights violator from a police line-up which may or may not contain the actual perpetrator. For each of these examples, it is the combination of uncertainty regarding the presence of the target with the task of identification that creates the more complex, and in some sense more naturalistic, decision environment.

1.1 The Compound Decision Environment

In these kind of compound decision environments, there are two separate decision problems that need to be considered. The fact that the target may or may not be present defines what is known as a detection problem (i.e., a target-present vs. target-absent discrimination). On the other hand, in those cases where the target happens to be present, picking out the target from among similar looking distractors is what is known as an identification problem (e.g., choosing the correct answer on a multiple choice test). Traditionally, research and theory development has focused on understanding these two decision problems in strict isolation (for many examples of such models see Egan, 1975; Macmillan & Creelman, 1991). The result has been twofold: A rich corpus of decision models aimed at understanding the pure forms of each decision problem; and, a dearth of models aimed at understanding complex decision tasks formed from a combination of both decision problems. This latter outcome is unfortunate because humans in the real world often find themselves confronted with compound decision tasks; containing both detection and identification components.

1.2 Decomposition of the Decision Task

The proper approach to understanding decisions made within a compound decision environment is to decompose the environment's complexity back into its constituent detection and identification components. Decomposition is necessary because performance in the kind of complex decision environments mentioned above will be determined either separately or conjointly by factors which affect the individual detection and identification components. Using a congealed metric (i.e., a composite variable comprising both detection and identification components) to measure performance in a compound decision task can lead to ambiguity regarding the locus of performance. To avoid this ambiguity, measuring performance in a compound decision environment would entail some method of separating out the constituent effects. The model presented in this report was developed with this kind of decomposition in mind.

2. Background on the Detection and Identification Components

The model described here has its roots in Signal Detection Theory (SDT), and consequently borrows many assumptions from that class of models (for an overview see Macmillan & Creelman, 1991). It is important to point out that SDT models are primarily decision models. They specify the rules and procedures for how decisions are made given certain kinds of information (defined in SDT as the decision variable). They generally do not specify, at least in any complicated way, how the decision variable is created. On the other hand, they do specify how information is used when making a decision. That said, this model should not be interpreted as a cognitive theory of the acquisition or storage of information pre se (i.e., where information used in the decision came from). However, because decision behaviour will be dependent on other cognitive processes such as memory storage, retrieval, recoding strategies, and decision criteria placement, parameter estimates from the SDT decision models can readily be used to make inferences about related cognitive processes or cognitive states relevant to the decision process (e.g., the robustness of information or factors that affect how the decision criteria are used).

In a very real sense, the proposed model is a compound decision model representing the fusion of a SDT detection model and a SDT identification model, each representing that particular aspect of the compound decision task. The model is a direct variant based on previous work in the area of uncertain identification and owes much to the modeling work of Starr, Metz, Lusted, and Goodenough (1975) and many others (e.g., Broadbent, 1958; Macmillan & Creelman, 1991; Nolte & Jaarsma, 1967; Swensson & Judy, 1981; Tanner & Norman, 1954). To better understand how these two aspects of the decision model come together, the following is a brief overview of the pure forms of detection and identification; each couched in terms of a formalized experimental procedure. Following each overview is a description of how the pure task can be modeled within the SDT framework. From these pure models it will be easier to show how a compound decision model can be constructed.

2.1 Description of Pure Detection Tasks

The following pure detection tasks are predicated on the idea of detecting the presence or absence of a target stimulus. Hence, there are two classes of stimuli, one representing some kind of target, the other representing the absence of the target. Although defining the absence of a target as a stimulus may seem contradictory, what is important is the relationship between the procedural environment in which the decision takes place and the decision task a person is asked to perform.

2.1.1 Simple Detection

In a simple detection task, participants are presented with a single stimulus that could be either a target or a foil, where the foil is a suitable non-target stimulus representing the absence of the target (see Egan, 1975; Macmillan & Creelman, 1991). In a simple detection procedure, only one of the stimuli is presented on any given detection trial. The goal is to detect the presence and absence of the target stimulus by properly categorizing the two classes of stimuli as they are presented. The extent to which the participant can properly categorize the stimuli reflects the participant's ability to discriminate between the two classes. In perception research, a foil stimulus might be white noise and a target a tone embedded in white noise. In memory research, target stimuli could be words that appeared on a previously presented list whereas foil stimuli would be any words not presented on the list.

2.1.2 1-of-*m* Detection

Consider a more complex version of the detection task where targets and foils are not presented as single stimuli, but multiple stimulus arrays (see Macmillan & Creelman, 1991). Now, a trial containing the "target" might consist of an array of several foil stimuli plus a single target stimulus. Likewise, the "foil" stimulus would be an array consisting entirely of foil stimuli. The goal is still the same as in simple detection, namely to detect which array contains the target. Note that despite the presence of multiple stimuli, the participant is not required to identify which stimulus is the target. They are merely required to discriminate an array containing the single target from an array that does not contain the target. This kind of detection task is often referred to as 1-of-*m* detection because the "target" array is composed of m possible stimuli, one of which may be a single target stimulus.

2.1.3 SDT Approach to Modeling Detection

The most basic assumption made for SDT modeling of a pure detection task is that a target stimulus will, on average, generate a numerically larger internal signal than a foil stimulus (Green & Swets, 1974). In SDT parlance, the internal signal generated in response to the stimulus is called the decision variable. Modeling the task requires a decision rule for how to treat the various internally generated signals. The decision rule for simple detection is simply a matter of establishing a criterion along the signal continuum. Any stimulus that exceeds the criterion is called a target, any that fall below are called foils.

In 1-of-*m* detection, the situation is more complicated because there are many stimuli (i.e., internal signals) to consider at once. One possible decision rule is to evaluate the signals from each stimulus individually (e.g., Swensson & Judy, 1981; Hacker & Ratcliff, 1979). If at least one signal exceeds the criterion a target present response is given. Another possible decision rule is to sum up all the individual signals. Presumably, the sum containing the target will be greater than a sum composed of all foils. If the sum exceeds the criterion, the response is to categorize the stimulus as a target.

2.2 Description of Pure Identification Tasks

Identification differs from detection primarily in the decision task assigned to the individual. Instead of detecting a target, the principle task is to uniquely specify the identity of a stimulus. In this sense, all the stimuli are targets, and the problem is to attach the proper identifying label to each stimulus.

2.2.1 Simple Identification

In a simple identification task participants are presented, as in simple detection, with a single stimulus from a predefined set. However, the task now is to identify the stimulus by specifying the category the stimulus belongs to. Such tasks are often used in perception experiments where the stimuli might be a set of tones of different frequencies; in which case the task is to identify the frequency of each tone (e.g., Tanner, 1956). In this kind of task, all of the stimuli are "targets" for identification and the difficulty is assigning the correct identity to the stimulus. The extent to which the participant can correctly identify which stimulus is which reflects their ability to discriminate between the stimuli.

2.2.2 *m*-Alternative Forced Choice (*m*AFC) Identification

Like 1-of-*m* detection, this task differs from simple identification in that there are more stimuli to consider before making a response (e.g., Hacker & Ratcliff 1979). For example, in a multiple choice test, all the options are presented simultaneously, and each option represents the answer to a question. The goal is to identify the answer to the particular question being asked. Note that, unlike 1-of-*m* detection, each option is a possible target in the sense that it is an answer to some question or other. This leads back to the important distinction between detection and identification tasks. In the latter, a target is always present (or assumed to be present). It would not make much sense to have a multiple choice exam in which the correct answer did not appear as one of the options. The idea that detection represents a "presence versus absence" discrimination, whereas identification represents an "always present" target identification is crucial to how the these two decision problems come together to form the compound decision task.

2.2.3 SDT Approach to Modeling Identification

The approach to identification is similar to detection in the sense that the decision rule is applied to the internal signals generated by the individual stimuli. For simple identification, an identification label is applied to the target signal that most closely matches the signal for that label. As with 1-of-*m* detection, the case of *m*AFC identification requires more than one signal to be considered (e.g., Macmillan & Creelman, 1991; Smith & Duncan, 2004). For multiple signals present with *m*AFC decision tasks, only one signal will presumably provide the closest match to the target to be identified.

As with detection, the response probabilities need to be transformed to represent the fundamental elements in the decision task (i.e., the individual signals). The importance of this transformation and how it is done will be covered in the next section.

3. Fusion of Detection and Identification: SDT Compound Decision (SDT-CD) Model

In both detection and identification, a representation of individual stimulus signals must be derived from the observed response probabilities (i.e., the proportion of correct responses made in each category). The transformation of response probabilities permits a reduction of the decision task down to its most basic elements (i.e., individual stimulus signals). Reduction is crucial because it provides a fundamental level of representation upon which the detection and identification components of the model can make contact with each other within the compound decision task. This level of representation is the point of fusion between detection and identification. That is, the point at which both components rely on the same fundamental source of information.

The actual form of the transformation depends on assumptions about the decision rule used in the compound decision task (Macmillan & Creelman, 1991). The following section describes the formal characteristics which define the compound decision task and the corresponding response probabilities associated with it. Subsequent sections describe various decision rules and how each rule defines a unique transformation of the response probabilities. Finally, the full model is presented.

3.1 Response Probabilities of the Compound Decision Task

A complete procedure for conducting experiments using a compound decision task which can be applied to the SDT-CD model utilizes both a target present (TP) and a target absent (TA)stimulus array as described above. The task of a subject in such an experiment is to either identify which stimulus is the target, or specify that the target is not present in the array (reject the array).[1] Assuming the procedure is done correctly, the participant is informed that the target may or may not be present, and so discriminating a TP versus TA array is a proper case of 1-of-*m* detection. On the other hand, identifying a stimulus within the array as the target is a *m*AFC identification task. Of course, from the decision maker's point of view the target is assumed to be present when a choice is made. It does not make much sense for one to make an identification response if they believe the target is not present.[2]

---

[1]In some cases the participant may be allowed to respond “don’t know” if they cannot make a determination one way or the other (see Wells, 1984). This response option is being omitted from the description for simplicity, but will be taken up again later.

[2]It is possible to require the participant to make an identification response even though they have specified the stimulus array as “target absent”. Such would be the case in tests of subliminal memory (see Macmillan & Creelman,

Table 1. Example Data

Response Choices	*TP Array*	*TA Array*
Picks Target	*HT* = .53	N/A[a]
Picks Foil	*FID* = .34	*FA* = .64
Reject	*MS* = .13	*CR* = .36

[a]Responses in this cell are not defined because the TA array contains only foils.

Table 1 summarizes the compound decision task and defines the various response probabilities that can be collected from such a task. Each cell in the table is defined by a descriptor variable and represents the possible response outcomes. Numerical values in each cell correspond to hypothetical data represented as proportions of responses accumulated over many successive trials. Also note that the values in each column of the table sum to 1.0. This represents the fact that the response options define the complete set of possible responses a participant can make regarding either a TP or TA array.

3.1.1 Response Outcomes for TP Arrays

When the participant correctly selects the target from a TP array, that response is defined as a correct identification with the proportion of such responses referred to as the correct ID rate or *CID*. When a foil is selected instead of the target, the response is called a false identification with the proportion of such responses referred to as the false ID rate or *FID*. Finally, deciding that the target is not present is called a miss with the proportion of such responses defined to the miss rate or *MS*. These response outcomes along with hypothetical values are given in the first column of Table 1.

3.1.2 Response Outcomes for TA Arrays

When the participant selects a foil from a TA array, the response is called a false alarm with the proportion of such responses referred to as the false alarm rate or *FA*. Correctly deciding that the target is not present is called a correct rejection with the proportion of such responses defined to the correct rejection rate or *CR*. These response

1991). The model is fully compatible with this kind of decision task, but examination of such procedures are beyond the scope of the present report. This issue will be addressed in the discussion.

outcomes along with hypothetical values are given in the second column of Table 1. Note that the top most cell is not defined for TA arrays because no target is present so a correct ID is not possible.[3]

3.2 Separation of Observed Response Probabilities

Before applying the SDT-CD model to data, the observed response probabilities from a compound decision task (see Table 1) must first be separated into probabilities representing the 1-of-*m* detection and *m*AFC identification components. These can then be further transformed into response probabilities corresponding to individual stimuli according to the decision rule assumed to be used by the decision maker. Some of the response probabilities corresponding to the detection or identification components can be taken directly from Table 1. Others must be derived.

3.2.1 Detection Components

The proportion of responses where a participant chooses to identify the target in a TP line-up represents the 1-of-*m* detection hit rate or *Hm*. From the values in Table 1, the 1-of-*m* detection hit rate is, $Hm = HT+FID = .53+.34 = .87$. Note that the 1-of-*m* detection hit rate is a sum of response probabilities because it represents the decision to choose a target, not whether the choice was correct or not. On the other hand, the proportion of responses where a participants chooses to identify the target in a TA line-up represents the 1-of-*m* detection false alarm rate or *Fm*. From Table 1, the 1-of-*m* false alarm rate is, $Fm = FA = .64$. Once computed, these 1-of-*m* detection probabilities can then be transformed into detection probabilities representing individual stimuli, where the transformation depends on the particular decision rule that is assumed to be in use for the detection component of the compound decision task.

3.2.2 Identification Components

Unlike the detection component, response probabilities for the identification component are derived from response to TP arrays only. For those participants who chose to identify the target in a TP array, the proportion who chose the target represents the *m*AFC hit rate,

[3]Some experimental procedures in the area of eyewitness memory use TA arrays which contain a mock target called an innocent suspect (e.g., Juslin, Olsson,& Winman, 1996). A mock target is used to simulate the real world case where the police have arrested an innocent suspect who could be falsely identified as the perpetrator. The use of TA arrays with mock targets can be handled in the SDT-CD model by including an additional parameter, but for simplicity is being omitted from the general model description.

*Hm*AFC. From the response probabilities in Table 1, the *m*AFC hit rate is, *Hm*AFC = *HT*/*Hm* = .53/.87 = .61. The proportion of those who chose a foil from a TP array represents the *m*AFC false alarm rate, *Fm*AFC. The *m*AFC false alarm rate is defined to be, *Fm*AFC = *FID*/*Hm* = .34/.87 = .39. Because the identification component only applies to responses where a choice is made from a TP array, the transformed proportions are made conditional on the probability of choosing by dividing the response probabilities from Table 1 by *Hm*. This conditionalization is an important transformation because it deconfounds identification from detection by separating out the influence of the detection probability from the identification probabilities. That is, identification is confined to choosing the target from among the foils when the target is present, regardless of how likely one is to want to make that choice in the first place (i.e., detection). Mathematically, conditionalizing causes the sum of *Hm*AFC and *Fm*AFC to be 1.0. As with the detection probabilities computed above, the values of *Hm*AFC and *Fm*AFC are applied to the model depending on the decision rule assumed to govern identification.

3.3 Decision Rules and the SDT-CD Model

The fusion of the models for detection and identification operate at the fundamental level of individual stimuli. Hence the decision rules for each component will also operate at the level of individual stimuli. With the 1-of-*m* detection and identification response probabilities extracted from the observed response probabilities (like those given in Table 1), they can be further transformed to represent response probabilities for individual stimuli according to the decision rule. Borrowing from previous SDT work (e.g., Starr et al, 1975; see also Macmillan & Creelman, 1991), two decision rules for detection and one for identification were explored in the development of the SDT-CD.

3.3.1 Decision Rules for Detection

For detection, the two decision rules are the independent observation rule and the integration rule. Both of these were adapted to the SDT-CD model from versions of SDT models developed for pure 1-of-*m* detection tasks. Once the 1-of-*m* detection response probabilities have been transformed to fit the decision rule, the standard SDT model for 1-of-*m* detection using that decision rule can used to fit the detection part of SDT-CD.

1. Independent Observation Rule

This is a well known decision rule in perception and memory research. Decision variables are assigned values independently for each stimulus in the array, then all decision variables are compared to a single criterion (Hacker & Ratcliff, 1979; Macmillan & Creelman, 1991; Starr et al, 1975). A "target present" response is made if at least one stimulus exceeds the criterion. Assuming response probabilities representing the detectability of individual stimuli could be measured, a simple detection model could be used to estimate parameters corresponding to the detectability of the target stimulus in the array. However, unless the detectablility of a single stimulus is directly measured beforehand, it must be inferred from 1-of-*m* detection probabilities. Fortunately, such inferences can be made given the constraints of the independent decision rule.

A false alarm occurs in 1-of-*m* detection when at least one of the m foils of a TA array exceeds the criterion. Given all possible combinations of foils falling above or below the criterion, *Fm* is equivalent to one minus the probability of all foils falling below the criterion. Consequently, the 1-of-*m* false alarm rate can be expressed as a function of the false alarm rate of an individual stimulus by,

$$Fm = 1 - (1 - F1)^m \tag{1}$$

where *F1* is the false alarm rate of an individual stimulus.[4] By rearranging Equation 1, an estimate for *F1* can be computed from *Fm* by,

$$F1 = 1 - (1 - Fm)^{\frac{1}{m}} \tag{2}$$

A hit in 1-of-*m* detection occurs when at least one of the stimuli in a TP array, the m-1 foils or the target, exceeds the criterion. Applying similar logic as before, this will be equal to one minus the probability of all m-1 foils and the target falling below the criterion. This is expressed as,

$$Hm = 1 - (1 - H1)(1 - F1)^{m-1} \tag{3}$$

where *H1* is the hit rate for the target. Rearranging to isolate *H1* and substituting Equation 2 for *F1* gives,

[4]Here it is assumed the false alarm rates of individual foil stimuli are all equal.

$$H1 = 1 - (1 - Fm)^{\frac{1-m}{m}} + Hm(1 - Fm)^{\frac{1-m}{m}} \tag{4}$$

The estimates for *F1* and *H1* represent the response probabilities for individual stimuli according to the independent observation rule. These response probabilities can now be fit with a standard simple detection SDT model. In other words, the detection component of SDT-CD is just a simple detection SDT model applied to response probabilities transformed from the compound decision task according to the independent observation rule.

The standard simple detection SDT model contains two parameters. The first, called *d'*, is a measure of how well a person can discriminate the presence of a target from foils. Larger values imply better discrimination. The other parameter called *c*, is a measure of response bias. Response bias represents a person's willingness to choose a target from the TP or TA array and is assigned a value of *d'*/2 when responding is unbiased (i.e., when there is no difference in preference between choosing and rejecting). The response probabilities *F1* and *H1* can be written as functions of *d'* and *c*,

$$H1 = \int_c^{\infty} f(x - d')dx \tag{5}$$

$$F1 = \int_c^{\infty} f(x)dx \tag{6}$$

where the function *f*(•) is the normal density function. Once the model is fit to the simple detection response probabilities, fitted probabilities can then be transformed into fitted observed response probabilities using Equations 1 and 3.

2. Integration Rule

This decision rule is similar to the independent observation rule in that values are assigned to each stimulus in the array. However, the actual decision variable is the sum of all the stimulus values, which is then compared to a criterion (see Graham, Kramer, & Yager, 1987; Tanner & Norman, 1954). If this sum exceeds the criterion, a target present response is made, otherwise a target absent response is given. Unlike the independent observation rule, no transformation of the 1-of-*m* detection probabilities is necessary for this rule. Consequently, *Hm* and *Fm* can be computed directly,

$$Hm = \int_{c}^{\infty} f\left(x - \frac{d'}{\sqrt{m}} \right) dx \qquad (7)$$

$$Fm = \int_{c}^{\infty} f(x) dx \qquad (8)$$

where *m* is the number of stimuli in the array.

3.3.2 Decision Rules for Identification

For identification, the decision rule adopted was an unbiased independent observation rule (e.g., Hacker & Judy, 1979; Starr et al, 1975). Specifically, it is assumed that people use a standard SDT *m*AFC decision rule to select a target once they have decided the target is present (i.e., using one of the detection decision rules described above). The general form of this rule is to choose the stimulus whose value of the decision variable most closely matches the intended identification target. In the model, the *m*AFC hit-rate is defined as the probability that similarity of the target stimulus to the intended identification target exceeds all of the foil stimuli across all possible values of the target stimulus. This unbiased probability, as a function of *d′*, can be written as,

$$HmAFC = \int f(x - d')z(x)^{m-1} dx \qquad (9)$$

where *z*(x) is the inverse normal function, *m* is the array size, and *d′* in this equation refers to the same *d′* as in Equation 7. Note that the limits of integration (not shown) run from negative to positive infinity because the probability of the target having a value larger than all foils must be computed for all possible values of the target decision variable.

3.4 Estimates Derived from Fitting the Model

Fitting the model to data produces model estimates for each of the observed response probabilities and two parameter estimates,

- *d′*, detectability of the target from the foils in identification and detection.
- *c*, willingness to choose a target from the stimulus array.

The extent to which the estimated response probabilities match the observed probabilities is a measure of how well the model fits the data. The parameter estimates are a summary explanation of what decision mechanisms were responsible for creating the data according to the model. In SDT-CD, the model explains the data in terms of the parameters d' and c.

It is important to consider the relationship between Equations 5, 7, and 9, in that d' in each equation refers to the same quantity, namely, the detectability of the target stimulus relative to the foils. This is the point where detection and identification fuse in the SDT-CD model because both decision components are assumed to depend on the same quantity, viz., the detectability of a target among foils. This is also why estimates of d' requires that response probabilities for detection and identification are fit simultaneously. This approach is where SDT-CD differs from other SDT models (e.g., Starr et al, 1975) and other process models of the compound decision task (e.g., Clark, 2003).

Because identification is assumed to be unbiased, parameter c does not occur in the identification equation. Hence, the bias parameter only refers to the detection component. However, the decision rule for identification is such that it could be assigned a separate response bias parameter. Doing so would not create problems for model validity. Although treating identification responses as biased would add an extra parameter to the model, it would also increase the number of free data points by an equal amount. Consequently, the fit of the model would not become over parameterized if such a change where implemented (for issues of over-parameterization see Pitt, Myung, & Zhang, 2002). However, the assumption that identification could be biased was not included in the present version of the SDT-CD model for several reasons. First, estimating bias for identification requires collection of considerably more data. In most experimental contexts, the practicality of such data collection is untenable for array sizes greater that two. Second, the interpretation of bias in identification is much different that detection. Identification bias is a propensity to favour a certain location in the array as containing the target. In some contexts, this may represent a variable of theoretical interest, but in most cases it is simply a nuisance variable.

4. Demonstrations

In this section the model is fit to several data sets representing various kinds of compound decision tasks. The purpose of these demonstrations is, (a) to show that the model can fit data from compound decision tasks, and (b) that parameter values derived from the model can be meaningfully interpreted with respect to experimental conditions. Actual choice of data sets was determined by a number of factors; including whether the experimental procedure used met the requirements for a full compound decision task, and whether it was possible to obtain access to the full data set. Given these requirements, choice was limited to only a handful of data sets, but these still managed to span across some diverse areas of research in compound decision tasks.

Because the chosen experiments were not conducted with the model in mind, the various experimental conditions are not necessarily the best choice with which to showcase the model's strengths or test assumptions of the model. In general, the model is not being used as a predictive model; at least in the sense that the parameter values for various experimental conditions are predicted a priori by the model. Instead, SDT-CD like most SDT decision models is a descriptive model. The value of descriptive models lies in changes to parameter values across conditions, providing a simplified interpretation of data from tasks where a direct analysis of the raw data would be a complicated affair; enough to cause misinterpretation of the results. I have argued that this is the case with compound decision tasks. Understanding the results from a direct analysis of the raw data can be difficult and tedious without the benefits of a proper model to simplify the results. The test of the model is grounded in how well it fits the data, and whether changes to parameter estimates across conditions make some theoretical sense.

Two separate models were created from a factorial combination of the two detection decision rules with the single identification decision rule. The two models were labelled Independent Observation and Integration referring to the varying detection decision component (note that the identification rule does not change). For each demonstration, the fit of both models was compared and the best fitting decision model chosen for any subsequent analysis of the data. Optimal parameters were found by non-linear optimization of the likelihood function compiled from Equations 5, 6, and 9 for the independent model, and Equations 7, 8, and 9 for the integration model (see Olgilvie & Creelman, 1968; Eliason, 1993; the exact equations used are given in Annex A). Optimization of the likelihood functions was done using the program Mathematica (Wolfram Research Inc., 2005). To ensure a good local minimum was reached, the optimization procedure was rerun several times on each data set using different parameter values. Goodness-of-fit was assessed using a χ^2 test.

Fitting the model to a data set results in estimates of each response probability and corresponding values of the two parameters. Because the goal is to obtain model estimates that are as close to the data as possible, one is actually hoping for a non-significant test result when comparing the data to model estimates. This is opposed to wanting a statistically significant result, as is usually the case in statistical analysis. Owing to the fact that the various experiments were not designed to test the model, one should be cautious of accepting a good fit. Consequently, a liberal rejection criteria of .10 was adopted for all χ^2 goodness-of-fit tests making it easier to reject the fit of the model. Both the Integration and Independent Observation decision models were fit to all data sets to determine which decision rule best fit the data. The implications for this will be taken up in the discussion section.

4.1 Tumour Detection in X-rays

This demonstration was conducted on data from an experiment by Starr et al (1975) designed to simulate the detection of tumours in chest X-ray radiographs. The study was originally conducted to test a compound decision model that is analytically identical to SDT-CD. The key difference between their model and SDT-CD is the assumption by Starr et al that detection and identification do not share the same d' parameter. In contrast, SDT-CD adopts the strong assumption that descriminability in detection and identification are derived from the same decision variable. In addition, the Starr et al model was limited to the Independent Observation rule. SDT-CD is very much an extension of the Starr et al model; albeit one that is more theoretically constrained but also more flexible to permit assessments of optimality in decision making. In their paper, the test of the Starr et al model was to predict identification Receiver Operating Characteristic (ROC) curves from detection ROC curves using an independent observation decision rule. The point of this demonstration is to use SDT-CD to test the difference between the two decision rules described above: Independent Observation and Integration.

For the experiment, sample radiographic images displaying the typical random mottling found in X-rays of chest tissue were created either with or without the faint image of a tumour. The presence of a tumour was simulated using a circular disk exposed onto the radiographic image. Each sample radiograph was divided into four quadrants, one quadrant of which may or may not have contained a faint image of a tumour. On each trial, subjects were shown a sample radiograph containing four quadrants and asked to either reject the image as tumour free, or indicate which quadrant contained a tumour. This procedure is therefore one in which all stimuli in an array are presented simultaneously. Because of experimental control over the creation of sample radiographs, it was possible to run each subject for approximately 100 trials per session over two sessions. The two sessions were conducted consecutively with a short break in between. Subjects were all Subject Matter Experts (SME) in the area of radiographic imaging and research consisting of 4 physicists and 1 technician.

For this demonstration, data were divided into separate sessions. Both model parameters, *d′* and *c*, were allowed to vary across sessions. The Integration model provided the best fit to the data, resulting in a non-significant deviation, $\chi^2(2) = 0.47$, $p < .80$. Fit of the Independent model, on the other hand, produced a significant deviation from the data, $\chi^2(2) = 7.88$, $p < .02$. Table 2 shows the data along with the fitted response probabilities produced by the Integration model. For brevity, only the *FA*, *HT*, and *CID* rates are shown.[5] An examination of Table 2 reveals that the response probabilities produced by the Integration model are very close to the observed data. The good fit of the Integration model implies two things. First, that detection and identification are yoked in the sense they are dependent on the same information (internally generated signal). Second, the decision rule which best describes the one used by subjects is the Integration rule.

Table 2. Results from the Tumour Detection Task.

	Session 1		***Session 2***	
Responses	***Data***	***Model***	***Data***	***Model***
FA	.33	.30	.30	.28
HT	.63	.65	.68	.69
CID	.50	.50	.58	.58
Parameters				
d′	1.81		2.12	
c	0.53		0.57	

[5]The remaining response probabilities can be computed from *FA*, *HT*, and *CID*, viz., *CR*=1-*FA*, *MS*=1-*HT*, and *FID*=*HT*(1-*CID*/*HT*). Hence, for this kind of compound decision experiment, there are only 3 degrees of freedom in the data.

Although there were no a priori predictions made for this demonstration, an examination of the parameter values reveals a few interesting observations. Because the subjects were all SMEs well practiced in reading radiographs, one would expect good performance with little change across trials. It is also reasonable to assume that these SMEs would adopt an optimal decision rule. Finally, given the dire implications for failing to detect a tumour, one might expect SMEs to adopt a fairly liberal criterion for deciding that a tumour is present. In keeping with the first of the expected results, values of d' are good for a detection memory task (e.g., 2.0-2.5, see Smith & Duncan, 2004) and there was only a slight change across sessions. As for the second expected outcome, the best fit of the Integration decision rule implies that these SMEs did not employ an optimal decision rule. Finally, the detection criterion for both sessions was below the value of $d'/2$, indicating the SMEs were liberal in their willingness to identify the presence of a tumour. However, the criterion was not as liberal as might be expected, especially considering the results from the next four demonstrations. It is possible the SMEs did not adopt a more liberal criterion because of the artificiality of the task.

4.2 Eyewitness Identification After Interrogation Stress

In a study on eyewitness memory, Morgan, Hazlett, Doran, Garrett, Hoyt, Thomas, Baranoski, and Southwick (2004) examined the effect of interrogation stress on the ability of subjects to later identify their abusers. The study was conducted in coordination with military survival school training offered to US military personnel. Relevant to the study was the part of the survival course involving wilderness evasion followed by mock captivity and interrogation in a prisoner of war camp. The rationale was to determine how stress would affect memory when tested in a naturalistic setting using a more applied memory test (eyewitness identification). Previous research has shown that memory for events perceived under stress can be good in controlled experimental settings (e.g., Canli, Zhao, Brewer, Gabrieli, & Cahill, 2000; Gold, 1992), but poor in naturalistic settings such as those experienced by military combat veterans (e.g., Roemer, Litz, Orsillo, Ehlich, & Friedman, 1998; Southwick, Morgan, Nicolaou, & Charney, 1997). However, because interrogation is a more perceptually focused experience free from many of the distractions found in combat, it was unclear how stress would affect eyewitness identification. For this demonstration, the two key interests were the effect of stress on both eyewitness performance and the decision rule used by subjects.

A total of 530 active-duty US military personnel enrolled in the survival course participated in the study. Subjects experienced both high and low-stress interrogations. High-stress involved actual physical confrontation of the trainee by the interrogator and a guard during the interrogation session. For the low-stress condition the physical confrontation was removed while maintaining all other aspects of the interrogation. Both the interrogator and the guard were unfamiliar to the subjects. The interrogation sessions lasted approximately 40 minutes and were conducted 4 hours apart. Order of interrogation stress was counter-balanced across subjects. In addition, all subjects experienced 48 hours of food and sleep deprivation before the interrogation sessions. Approximately 24 hours after the last interrogation session, subjects were released from the their mock captivity and given access to food and water. The eyewitness test was then administered in one of the school classrooms using a live line-up of survival school instructors acting as foils. The TP line-up included the subjects' interrogator. Like Demonstration 1, this experiment employed a simultaneous presentation procedure. There was only one presentation trial so each subject supplied a single response.

To test the effect of stress, both model parameters were allowed to vary across each interrogation condition. Models for both decision rules showed significant deviations from the data, with the Integration model producing a far better fit, $\chi^2(2) = 11.7$, $p < .01$, than the Independent model, $\chi^2(2) = 40.5$, $p < .001$. Table 3 shows the results of the study and the fit of the Integration model. Although the goodness-of-fit test for the Integration model was significant, the main point of error in predicted response proportions occurs in the *FA* rates. Otherwise, the model produced a very good fit to the *HT* and *CID* rates.

Table 3. Results from the Interrogation Stress Eyewitness Task.

	High Stress		***Low Stress***	
Responses	***Data***	***Model***	***Data***	***Model***
FA	.45	.59	.50	.65
HT	.85	.71	.95	.83
CID	.26	.25	.62	.57
Parameters				
d'	1.26		2.24	
c	-.023		-.039	

Part of the reason for the worse fit is likely do to the experimental procedure used for the eyewitness test. The decision rules for both models assumes that all foils are all equally discriminable from the target. However, because the line-up foils were not standardized before testing it is possible this assumption of the model was violated in the experimental procedure. Despite this the model results are quite striking. Performance in identifying the abuser in the low stress condition was almost twice as good as in the high stress condition. This result is consistent with the results of laboratory research despite the study being conducted in a naturalistic setting. Although Morgan et al (2004) reached a similar conclusion concerning better performance in the low stress condition, it was not so obvious from their onerous multiple comparisons of the various response proportions. The simple summary analysis provided by the model, that of reducing the complex data set down to a single value representing performance provided a clearer picture of the data. In addition, one important aspect of the experiment captured by the model was not mentioned in the analysis of Morgan et al. Namely, the fact that both groups showed a strong willingness to choose a perpetrator from the line-up (i.e., both adopted a very liberal response criterion). Although the model is over-estimating the *FA* rate of subjects, a strong bias to choose is still plainly evident (c.f. Demonstration 1). The failure of Morgan et al to mention response bias is not so surprising given that conventional analysis of response proportions do not provide an easy means of obtaining explicit measures of willingness to choose. It is unclear why subjects chose such a liberal response criterion, but it may have reflected the training atmosphere in the school, or pressure to be seen by the command hierarchy as producing a positive response.

4.3 Verbal Overshadowing Effect in Eyewitness Memory

Previous research in eyewitness memory has shown that performance in eyewitness tasks can be very low. Techniques to improve memory, such as producing a verbal sketch of the perpetrator in an eyewitness event, have been found to influence subsequent eyewitness performance. The effect on eyewitness memory of verbally re-describing (or recoding) an eyewitness event is called the Verbal Overshadowing Effect (VOE). Because the area of eyewitness research has lacked a coherent model of the eyewitness task (such as SDT-CD), debate in the literature has focused on performance vs. response bias as the locus of effect for the VOE. Because performance and bias are two key aspects de-confounded in the SDT-CD model, this data set is uniquely suited to an analysis by the model.

The data in this demonstration came from an experiment by Clare and Lewandowsky (2004) who manipulated the type of recoding subjects engaged in after viewing a staged crime event. Subjects were assigned to one of three conditions, two of which involved re-remembering the event by focusing on describing the global pictorial features (Holistic) or creating a point-by-point verbal description (Verbal). The Verbal condition was the standard VOE manipulation which has previously shown effects on eyewitness memory. The Holistic condition represented a way of re-remembering designed to minimize the impact of VOE on eyewitness memory. Approximately 2 hours after witnessing the staged event, subjects in the Verbal and Holistic conditions were given instructions to re-remember the event followed by an eyewitness test consisting of either a TP or TA array of size 6. Subjects in the control condition were not given any specific instructions for re-remembering the event, but instead were given the eyewitness test after performing 2 hours of unrelated activities. Like the previous demonstration, each subject was shown a single stimulus array (TA or TP) and hence provided only a single response in the experiment.

The approach to modeling taken in this demonstration was to create two versions of the SDT-CD model each representing one of the hypotheses to be tested; performance and bias. This was done by constraining parameter values across the three conditions. The criterion hypothesis was represented by an SDT-CD model in which d' was held constant but c was allowed to vary across conditions. The performance hypothesis was represented by a model in which d' varied but c was held constant. If the data correspond to one or the other hypotheses, then one model should fit while the other model is rejected. Before testing each model hypothesis, an initial fit of the unconstrained Independent and Integration models was done to determine which decision rule was best supported by the data. For these unconstrained fits, both parameters were allowed to vary across all three conditions.

Table 4. Results from the Criterion Hypothesis Model.[a]

	Control		***Holistic***		***Verbal***	
Responses	***Data***	***Model***	***Data***	***Model***	***Data***	***Model***
FA	.77	.75	.48	.47	.48	.40
HT	.93	.95	.81	.82	.64	.74
CID	.80	.80	.69	.70	.57	.63
Parameters						
d'	2.42		2.42		2.42	
c	-0.68		0.08		0.33	

[a]For this model, d' was held constant while c was allowed to vary across conditions.

The initial goodness-of-fit test resulted in a good fit of the Integration model, $\chi^2(3) = 5.38$, $p < .15$, but a significantly poor fit of the Independent model, $\chi^2(3) = 119.47$, $p < .001$. Given the initial fit results, models for each hypothesis were fit using the Integration decision rule. The outcome of the hypothesis tests were clear. The model representing the criterion hypothesis showed a good fit to the data, $\chi^2(4) = 5.45$, $p < .37$, while the model representing the performance hypothesis was rejected, $\chi^2(4) = 19.12$, $p < .001$. Data and model fit for the criterion hypothesis model are shown in Table 6. The obvious conclusion from these tests is that VOE causes a shift in response criterion but has no appreciable effect on performance.

In some sense this is a surprising result for it implies that recoding (or re-remembering) by verbally describing a previous eyewitness event has little effect on performance. Instead, it appears to profoundly effect response bias. From Table 6, subjects in the Verbal (VOE) condition were the most conservative in their willingness to select the perpetrator from the line-up. The Holistic condition also made subjects more conservative, but not quite as much as the Verbal group. The conservative bias of these two groups was in stark contrast to the control group who showed a strong willingness to identify the perpetrator; much like that seen in Demonstration 2.

4.4 Search and Rescue

This final demonstration utilizes data from a Search and Rescue (SAR) experiment conducted by Stager and Hameluck (1986). The purpose of their experiment was to derive a normative set of identification probabilities for modeling real life SAR by simulating a SAR operation using static naturalistic images. In the SAR task, they used actual aerial photographs of various terrain taken from two different altitudes (500 and 1,000 ft) and two different visual angles (vertical and 45º oblique) forming a 2×2 factorial set of images. A simulated crash site was painted within one of six randomly selected quadrants to create TP stimuli. Images representing TA stimuli were left unchanged. The photographs were projected on a screen approximately 5ft square so as to fill 40º of the subject's viewing angle. Like the tumour detection task of Demonstration 1, all subjects were SMEs in SAR and were run through approximately 100 trials of TA and TP stimulus arrays. On each trial SMEs were required to identify the location of a crash site or reject the image as not containing one. Like all previous demonstrations, the stimulus array was presented simultaneously.

Although the purpose of this experiment was not to compare altitude and visual angle, it seemed appropriate to compare these two factors as they were the only variables that were manipulated. The Independent and Integration models were fit to data set allowing both parameters to vary across all four conditions.[6] Unlike the previous demonstrations, the Independent model produced a good fit, $\chi^2(4) = 6.91$, $p < .15$, with the fit of the Integration model rejected as significantly deviating from the data, $\chi^2(4) = 14.2$, $p < .007$. Given this result, all subsequent model fits used the Independent decision rule model.

As in Demonstration 3, separate reduced models representing performance and criterion hypotheses were created to assess which of these two characteristics (or both) were affected by each factor (altitude and angle). For altitude, neither the performance hypothesis model, $\chi^2(3) = 0.51$, $p < .90$, nor the criterion hypothesis model, $\chi^2(3) = 0.62$, $p < .89$, could be rejected as both showed good fits to the data. A reduced model was fit to the altitude data producing a good fit when both parameters were held constant, $\chi^2(3) = 0.61$, $p < .96$. In contrast, fitting each hypothesis model to data in the angle condition produced a different pattern of results. The criterion hypothesis model was rejected, $\chi^2(3) = 7.45$, $p < .02$, whereas the performance hypothesis was not, $\chi^2(3) = 1.93$, $p < .60$. The data for each condition and model fits are shown in Table 6.

Table 5. Results from the SAR Task.

	Altitude[a]				***Angle***[b]			
	500 ft.		***1,000 ft.***		***Oblique***		***Vertical***	
Response	***Data***	***Model***	***Data***	***Model***	***Data***	***Model***	***Data***	***Model***
FA	.31	.34	.32	.34	.37	.34	.27	.34
HT	.85	.85	.86	.85	.82	.82	.89	.88
CID	.72	.71	.67	.71	.66	.65	.74	.76
Parameters								
d′	2.30		2.30		2.15		2.46	
c	1.50		1.50		1.50		1.50	

[a]Reduced model with both *d′* and *c* constant across conditions.

[b]Reduced model with only *c* constant across conditions.

[6]The four conditions created by the 2×2 factorial design where treated as separate.

The results of the model fitting suggest that altitude has little effect on identification of a crash site, at least in the context of this simulated SAR task. On the other hand, viewing angle appears to affect performance in identifying a crash site (vertical is better than oblique), but has little effect on willingness to choose a target. The fact that the Independent Observation model showed the best fit implies the SMEs in this experiment were using an optimal decision rule for the task. This is perhaps not so surprising as these particular SMEs were trained to scan each quadrant of the target area in a systematic fashion (McFadden, personal communication, November 24, 2006). This type of training would lead to separate assessments of each stimulus in the array; which is consistent with the assumptions of the Independent Observation rule.

5. Discussion and Conclusions

This section provides a summary of the model fits to the demonstration data sets, as well as a discussion of the model's advantages and other issues related to modeling compound decision tasks. A comparison of SDT-CD to other types of models and possible future directions are also discussed.

5.1 Summary of Model Fits

- Overall, the model fit 3 out of 4 data sets very well. Two of the data sets (Demonstrations 3 and 4) allowed for testing of models representing different theoretical hypothesis. The hypothesis tests illustrate the diversity of using SDT-CD as a statistical analysis tool to help better understand data. In the case of data from Demonstration 2 (eyewitness identification after interrogation stress), two possible reasons for the poor fit are that the SDT-CD model is simply wrong, or the experimental procedure violated assumptions of the model. Although the former is certainly a possibility, the fact that line-up foils were not pre-screened to ensure homogenous similarity to the target makes it impossible to know with certainty which explanation is correct. For comparison, foils used in the eyewitness line-ups of Demonstration 3, where the model showed a good fit, were pre-screened for similarity to the target.

- The results from the demonstrations also revealed an interesting trend in subject's choice of decision rule. In 3 out of 4 data sets, the Integration decision rule produced the best fit. Only the data from Demonstration 4 showed a superior fit of the Independent Observation rule. This has implications for decision making because the model decision rules do not differ merely on the basis of strategic approach to the task. When decision making is expressed in terms of optimal use of information for maximizing the discrimination of foils from target, the two decision rules presented here occupy opposite ends of the spectrum. The Independent Observation rule is by far the more optimal decision rule; being nearly equivalent to the analytically defined maximum likelihood optimum for the SDT-CD model (Graham, et al 1987). The Integration rule, on the other hand, is far from optimal. Although the sample of data sets is small, the pattern is quite clear. People are more likely to employ a non-optimal decision strategy when faced with a compound decision task. This conclusion is supported by a recent fit of SDT-CD to several data sets exclusively within the eyewitness domain in which a vast majority were best fit by the Integration rule (Duncan, 2006).

- The decision rule which best fit the data in each demonstration is also related to variations in how the task was approached by subjects. The three data sets that were best fit by the Integration rule all used subjects and SMEs with no explicit training on how they should perform the task. The data from Demonstration 4 however, employed SMEs who where given explicit training to assess the array stimuli in a manner consistent with the optimal decision rule. This implies two conclusions. First, that the model can be used to assess whether subjects are using an optimal decision rule. Second, that training in procedures consistent with the optimal decision rule lead to optimal decision making. There is also a suggestion that procedural changes can cause use of the optimal decision rule. In the area of eyewitness identification, it has been argued that sequential presentation (array stimuli presented 1 at a time instead of simultaneously) encourages the use of the optimal decision rule by discouraging global assessment strategies like the Integration rule (e.g., Clark & Davey, 2005; McQuiston-Surrett, Malpass, & Tredoux, 2006; Steblay, Dysart, Fulero, & Lindsay, 2001). Although not explicitly tested here, SDT-CD can also be used to assess this claim.

5.2 Advantages of the Model

There are three distinct advantages to using SDT-CD for analysis of data from compound decision tasks like those described here. For the most part, the advantages of SDT-CD derives from the transformation of the complex interplay of response probabilities into parameter values representing simple measures of performance and response bias.

5.2.1 Simplification of Data Analysis: Separation of Performance from Bias

First, and possibly foremost, is the simplification of data analysis and interpretation. The standard compound decision task generates three separate sets of response probabilities, *FA*/*CR*, *HT*/*MS*, and *CID*/*FID*. A more traditional analysis would be to conduct separate comparisons for each set. For an experiment with only two conditions, this type of analysis would produce three distinct significant or non-significant outcomes (from a total of eight possible). With three experimental conditions, one would be faced with interpreting the pattern of significance from as many as nine pairwise comparisons. Clearly, the complexity of a conventional the data analysis grows geometrically with number of experimental conditions. Even with only three, one is faced with an overwhelming pattern of test results from which to derive effects on performance. In contrast, a three condition experiment fit with SDT-CD would require at most three comparisons. A two condition experiment requires only one.

To compound the complexity problem, it is not obvious from a conventional analysis whether significant differences in response probabilities reflect changes in performance, changes in response bias, or both. Because SDT-CD deconfounds the effects of performance from response bias, interpretation of the data is not only simplified, but also much clearer as to how experimental conditions affect decision making.

5.2.2 Tests for Optimal Decision Making

Another key advantage of SDT-CD is the ability to test whether people are using an optimal decision rule. As shown in the demonstrations, both the Integration and Independent decision models do not always fit equally well, so some assessment of optimal rule use can be made. Unlike conventional analysis of response probabilities which provide no information about rule use per se, tests for optimal decision making are trivial to obtain using SDT-CD.

5.2.3 Explicit Hypothesis Testing

As shown in Demonstrations 3 and 4, SDT-CD can be used to test specific hypotheses regarding performance and response bias. The general approach is to use theory to constrain parameter values across experimental conditions and test the fit of the model. In a statistical sense, SDT-CD takes on the role of the traditional Null Hypothesis. Rejecting the fit of the model amounts to rejecting the Null Hypothesis. This approach to models is by no means a unique property of SDT-CD, and has been widely employed as a general approach in many areas of cognitive modeling (for examples see Pitt et al, 2002).

5.3 Extensions of SDT-CD

The SDT-CD model is adaptable to variations of the standard compound decision task. Given its SDT roots, the model can generally be adapted to variations open to other SDT-like models.

5.3.1 Extra Response Options: "Don't Know" and Confidence

Previously it was mentioned that some experimental procedures, particularly in the field of eyewitness identification, have added a "don't know" response option. Other procedures require subjects to make confidence ratings about their decision. The addition of response options and confidence ratings are typically handled within a SDT framework by adding extra criteria to the model. These extra criteria are used to divide each confidence category in the same way the single detection criteria is used to divide the *FA* and *HT* rates. In the case of SDT-CD, the same approach can be used to easily implement "don't know" and confidence ratings into the model. SDT, and by implication SDT-CD, treat the former as just a special case of the latter. That is, "don't know" responses are treated as low confidence responses. Although not presented as a demonstration, data from Wells, Rydell, and Seelau (1993) was well fit by a SDT-CD model extended to include the addition of "don't know" responses that was added to their procedure. An example of modeling confidence can be found in Starr et al (1987), who successfully fit a version of SDT-CD to confidence ratings data.

5.3.2 Calibration

The ability to handle confidence ratings makes SDT-CD immediately applicable as a model for examining confidence calibration (i.e., the confidence/accuracy relation, see Baranski & Petrusic, 1994; Brewer, Keast, & Rishworth, 2002; Olsson, Juslin, & Winman, 1998; Wagenaar, 1988).[7] The SDT approach to calibration has been theoretically formalized by Gu and Walsten (2001), and SDT-CD borrows directly from this work. To model calibration, traditional confidence categories are assigned probability values instead of confidence labels (e.g., 80% instead of "very sure"). As in ratings data, criteria are used to separate the calibration categories and then fit to the model in the same manner as confidence ratings. The key importance of SDT-CD as a calibration model is that all of the advantages of modeling compound decision tasks can be readily extended to the study of calibration in these decision tasks as well (e.g., separation of performance and bias, simplicity of analysis etc).

---

[7]A distinction is made here between probability calibration (e.g., Ferrell, & McGoey, 1988) and confidence calibration (e.g., Baranski & Petrusic, 1994). The former refers to the relationship between subjective probability estimates and objective probabilities. Although procedurally there is often little difference between the two in experimental settings, the approach to modeling calibration in SDT-CD is equivalent to how the model handles confidence ratings. This implies a tighter theoretical binding between subjective probability and confidence that is not implied in probability calibration research.

5.3.3 Subliminal Effects on Decision Making

Presumably there may exist stimuli in the decision environment which can have an unconscious affect on decision making. These would be stimuli of which the decision maker is not only unaware, but which they cannot become consciously aware. The effects of such stimuli, should they be presumed to exist, can be examined using SDT-CD without changes to the model itself. The primary change occurs experimentally, by presenting a subliminal stimulus on some trials and requiring subjects to provide an additional identification response on all trials where they have indicated that no target is present. Analysis of such data can be done by restricting analysis of responses to subliminal TP presentations.

6. Recommendations

- Scientist should be encouraged to employ more naturalistic compound decision tasks in research settings. Barriers that may have prevented the use of compound decision tasks, such as complexity of data analysis, can be overcome with the use of SDT-CD.

- Researchers using compound decision tasks should be cognizant of ensuring the full standard procedure is represented in the task environment (e.g., use of both TP and TA arrays). This will permit the use of SDT-CD in data analysis, by those conducting the research or by others who may have an interest.

- Decision makers should try to encourage use of the optimal decision rule by training SMEs to individually assess each array stimulus (i.e., SAR scanning), or by forcing array stimuli to be presented in sequence. Adopting these constraints should lead decision makers to produce separate independent assessments of array stimuli consistent with the optimal decision rule.

- Training for decision making in compound decision tasks should emphasize scanning techniques or use sequential presentation to encourage adoption of the optimal decision rule for this kind of task.

- The model represents a formal abstract conceptual framework, and so should be applied to research in automated decision support systems.

7. References

1. Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics*, *55*, 412-428.

2. Brewer, N., Keast, A., & Rishworth, A. (2002). The confidence-accuracy relationship in eyewitness identification: The effects of reflection and disconfirmation on correlation and calibration. *Journal of Experimental Psychology: Applied*, *8*(1), 44-56.

3. Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon Press.

4. Canli, T., Zhao, Z., Brewer, J., Gabrieli, J. D. E., & Cahill, L. (2000). Event-related activation in the human amygdala associates with later memory for individual emotional response. *Journal of Neuroscience*, *20*(19), RC99.

5. Clare, J., & Lewandowsky, S. (2004). Verbalizing Facial Memory: Criterion Effects in Verbal Overshadowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(4), 739-755.

6. Clark, S. E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology*, *17*(6), 629-654.

7. Clark, S. E., & Davey, S. L. (2005). The Target-to-Foils Shift in Simultaneous and Sequential Lineups. *Law & Human Behavior*, *29*(2), 151-172.

8. Duncan, M. J. (2006). A signal detection theory model of eyewitness identification. Manuscript in preparation.

9. Egan, J. P. (1975). *Signal detection theory and ROC analysis*. New York: Academic Press.

10. Eliason, S. R. (1993). *Maximum likelihood estimation: Logic and practice*. Newbury Park: Sage Publications, Inc.

11. Ferrell, W. R., & McGoey, P. J. (1980). A model of calibration for subjective probabilities. *Organizational Behavior & Human Decision Processes*, *26*(1), 32-53.

12. Gold, P. E. (1992). *Modulation of memory processing: Enhancement of memory in rodents and humans*. In L. R. Squire & N. Butters (Eds.), Neuropsychology of memory (2nd ed.). (pp. 402-414): Guilford Press.

13. Graham, N., Kramer, P., & Yager, D. (1987). Signal-detection models for multidimensional stimuli: Probability distributions and combination rules. *Journal of Mathematical Psychology*, *31*(4), 366-409.

14. Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics*. Huntington, NY: Robert E. Krieger.

15. Gu, H., & Wallsten, T. S. (2001). On setting response criteria for calibrated subjective probability estimates. *Journal of Mathematical Psychology*, *45*(4), 551-563.

16. Hacker, M. J., & Ratcliff, R. (1979). A revisted table of d' for M-alternative forced choice. *Perception & Psychophysics*, *26*(2), 168-170.

17. Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence-accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1304-1316.

18. Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York, NY: Cambridge University Press.

19. McQuiston-Surrett, D., Malpass, R. S., & Tredoux, C. G. (2006). Sequential vs. Simultaneous Lineups: A Review of Methods, Data, and Theory. *Psychology, Public Policy, and Law*, *12*(2), 137-169.

20. Morgan, C. A. I., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., Baranoski, M., & Southwick, S. M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry*, *27*(3), 265-279.

21. Nolte, L. W., & Jaarsma, D. (1967). More on the detection of M orthogonal signals. Journal of the Acoustical Society of America, 41, 231-259.

22. Olgilvie, J. C., & Creelman, C. D. (1968). Maximun-likelihood estimation of receiver operating characteristic curve paramters. *Journal of mathematical psychology*, *5*(1), 377-391.

23. Olsson, N., Juslin, P., & Winman, A. (1998). Realism of confidence in earwitness versus eyewitness identification. *Journal of Experimental Psychology: Applied*, *4*(2), 101-118.

24. Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, *109*(3), 472-491.

25. Roemer, L., Litz, B. T., Orsillo, S. M., Ehlich, P. J., & Friedman, M. J. (1998). Increases in retrospective accounts of war-zone exposure over time: The role of PTSD symptom severity. *Journal of Traumatic Stress*, *11*(3), 597-605.

26. Smith, D. G., & Duncan, M. J. J. (2004). Testing Theories of Recognition Memory by Predicting Performance Across Paradigms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 615-625.

27. Southwick, S. M., Morgan, C. A. I., Nicolaou, A. L., & Charney, D. S. (1997). Consistency of memory for combat-related traumatic events in veterans of Operation Desert Storm. *American Journal of Psychiatry*, *154*(2), 173-177.

28. Stager, P., & Hameluck, D. (1986). Visual detection probabilities for search and rescue (Final report DSSC01SE9771137527). Toronto: Defence Research and Development Canada (formerly Defence and Civil Institute of Environmental Medicine).

29. Starr, S. J., Metz, C. E., Lusted, L. B., & Goodenough, D. J. (1975). Visual detection and localization of radiographic images. *Radiology*, *116*(3), 533-538.

30. Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, *25*(5), 459-473.

31. Swensson, R. G., & Judy, P. F. (1981). Detection of noisy visual targets: Models for the effects of spatial uncertainty and signal-to-noise ratio. *Perception & Psychophysics*, *29*(6), 521-534.

32. Tanner, W. P. (1956). Theory of recognition. *Journal of the Acoustical Society of America*, *28*, 882-888.

33. Tanner, W. P., & Norman, G. Z. (1954). The human use of information. II. Signal detection for an unknown signal parameter. *Transactions of the IRE Professional Group on Information Theory*, *PGIT-4*, 222-226.

34. Wagenaar, W. A. (1988). Calibration and the effects of knowledge and reconstruction in retrieval from memory. *Cognition*, *28*(3), 277-296.

35. Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology*, *14*(2), 89-103.

36. Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology*, *78*(5), 835-844.

37. Wolfram Research Inc. (2005). Mathematica (Version 5.2) [Computer software]. Champaign, IL: Wolfram Research Inc.

Annexes

The likelihood functions for each version of SDT-CD were derived from those by Olgilvie and Creelman (1968). For computational efficiency, the likelihood function was transformed into a sum of log-likelihoods.[8] The output of this equation is referred to as the Log Maximum Likelihood Estimator (LMLE). The following log-likelihood equation was used for both decision models, with the corresponding probabilities computed depending on the decision rule.

$$\begin{aligned} LMLE = {} & \ln\left((1-P(FA \mid TA))n_{CR}\right) + \ln\left(P(FA \mid TA)n_{FA}\right) + \ln\left(P(HT \mid TP)n_{HT}\right) + \\ & \ln\left((1-P(HT \mid TP))n_{MS}\right) + \ln\left((1-P(FID \mid TP))n_{FID}\right) + \ln\left(P(CID \mid TP)n_{CID}\right) \end{aligned} \quad \text{(A1)}$$

Where n is the response frequency for each respective response category, $P(FA|\mathrm{TA})$ and $P(HT|\mathrm{TP})$ are defined by equations 5 and 6 (for the Independent Observation rule) or equations 7 and 8 (for the Integration rule), and $P(CID|\mathrm{TP})$ is defined as in Equation 9. The parameters d and c are assigned values by the optimization function to maximize the value of Equation A1 (or equivalently to minimize its negative). Because Equation A1 does not have an analytic solution, the corresponding Fisher Information Matrix, and hence estimates of variance in the fitted parameters, can not be calculated using this model. Despite this, statistical tests on parameter values can still be possible by fitting the model to data from individual subjects.

[8] Likelihood functions are typically defined in terms of the products of likelihoods. Taking the natural logarithm transforms the product calculation into a summation which is more numerically efficient to calculate.

List of symbols/abbreviations/acronyms/initialisms

DND	Department of National Defence
d′	Model parameter representing decision making performance
c	Model parameter representing decision bias
χ^2	Chi-squared test statistic for assessing model goodness-of-fit
SDT-CD	Signal Detection Theory – Compound Decision model
TP	Array of stimuli containing foils plus the target
TA	Array of stimuli containing only foils
HT	Detection hit-rate defined as a response proportion
FA	Detection false-alarm rate defined as a response proportion
CID	Correct identification rate defined as a response proportion
FID	False identification rate defined as a response proportion
MS	Detection miss rate defined as a response proportion
CR	Detection correct-rejection rate defined as a response proportion
ID	Identification

Glossary

Technical term	Explanation of term
Target	Stimulus acting as the target to be identified
Foil	Stimulus acting as a non-target distractor
Hit rate	Proportion of correct target detections
False-alarm rate	Proportion of incorrect target detections
Correct ID rate	Proportion of correct target identifications
False ID rate	Proportion of incorrect target identifications
Miss rate	Proportion of targets not detected
Correct-rejection rate	Proportion of foils correctly rejected as non targets

Distribution list

Document No.: DRDC Toronto TR 2006-256

HARD COPY LIST: Internal Distribution by Centre:

2 DRDC Toronto Library
15 For the Author
1 David Smith: David.Smith@drdc-rddc-gc.ca
1 David Bryant: David.Bryant@drdc-rddc-gc.ca
1 Renee Chow: Renee.Chow@drdc-rddc-gc.ca
1 Wenbi Wang: Wenbi.Wang@drdc-rddc-gc.ca
1 Jocelyn Keiller: Jocelyn.Keillor@drdc-rddc.gc.ca
1 Sharon McFadden: Sharon.McFadden@drdc-rddc.gc.ca

23 TOTAL LIST PART 1

ELECTRONIC COPY LIST: External Distribution by DRDKIM

Adel Guitouni: Adel.Guitouni@drdc-rddc-gc.ca
Kendall Wheaton: Kendall.Wheaton@drdc-rddc-gc.ca
Walter Dyck: Walter.Dyck@drdc-rddc-gc.ca
Anne-Claire Boury-Brisset: Anne-Claire.Boury-Brisset@drdc-rddc-gc.ca
Abdeslem Boukhtouta: Abdeslem.Boukhtouta@drdc-rddc-gc.ca
Frederick Lichacz: Frederick.Lichacz@drdc-rddc-gc.ca
Philip Farrell: Philip.Farrell@drdc-rddc-gc.ca
Jacqui Creebolder: Jacqui.Creebolder@drdc-rddc-gc.ca
Francois Bernier: Francois.Bernier@drdc-rddc-gc.ca
Denis Gouin: Denis.Gouin@drdc-rddc-gc.ca
Micheline Belanger: Micheline.Belanger@drdc-rddc-gc.ca
Richard Breton: Richard.Breton@drdc-rddc-gc.ca
Valdur Pille: Valdur.Pille@drdc-rddc-gc.ca
Regine Lecoq: Regine.Lecoq@drdc-rddc-gc.ca

DOCUMENT CONTROL DATA

(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document, Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's document, or tasking agency, are entered in section 8.)

Publishing: DRDC Toronto

Performing: DRDC Toronto

Monitoring:

Contracting:

2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.)

UNCLASSIFIED

3. TITLE (The complete document title as indicated on the title page. Its classification is indicated by the appropriate abbreviation (S, C, R, or U) in parenthesis at the end of the title)

A signal detection model of compound decision tasks (U)
(U)

4. AUTHORS (First name, middle initial and last name. If military, show rank, e.g. Maj. John E. Doe.)

Matthew J. Duncan

5. DATE OF PUBLICATION (Month and year of publication of document.)

December 2006

6a NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)

56

6b. NO. OF REFS (Total cited in document.)

37

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of document, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Report

8. SPONSORING ACTIVITY (The names of the department project office or laboratory sponsoring the research and development − include address.)

Sponsoring:

Tasking:

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant under which the document was written. Please specify whether project or grant.)

16kf

9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document)

DRDC Toronto 2006−256

10b. OTHER DOCUMENT NO(s). (Any other numbers under which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (Any limitations on the dissemination of the document, other than those imposed by security classification.)

Unlimited distribution

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11), However, when further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited announcement

DOCUMENT CONTROL DATA

(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

(U) Detection and identification represent two fundamental types of decision tasks. Although research has focused on each in isolation, the pure forms of these tasks are generally not representative of more complex naturalistic decision environments. For example, a decision maker involved in a Search and Rescue (SAR) operation is faced with locating and identifying a crash site. This kind of decision environment is characterized by both detection and identification components. Namely, uncertainty regarding the presence of a target crash site, and the task of identifying the target from among similar looking structures in the terrain. Decision research using compound decision tasks (detection plus identification) has the advantage of making greater contact with naturalistic environments, but carries with it the cost of increased complexity in analyzing and understanding the data. Because compound decision tasks have more than one locus where decision making can be affected, a formal method is needed to disambiguate (deconfound) effects on decision making and simplify an understanding of decision making performance in complex tasks. In this report a formal model of compound decision tasks (SDT−CD) is presented which fulfills this role. The model was assessed by an analysis of several demonstration data sets from a wide variety of content domains which highlight its ability to simplify the complexity of the task and provide readily interpretable results. In addition to measures of performance and decision bias, the model can be used to test hypotheses about decision making and permits an assessment of whether decision making is optimal.

(U) La détection et l’identification représentent deux types fondamentaux de tâches décisionnelles. Ces deux tâches, prises isolément, ont fait l’objet de travaux de recherche. Cependant, dans leur forme pure, elles ne sont pas représentatives, en général, des environnements décisionnels du monde réel, beaucoup plus complexes. Par exemple, un décideur qui participe à une opération de recherche et de sauvetage (SAR) est chargé de localiser et d’identifier le site de l’écrasement d’un aéronef. Ce type d’environnement décisionnel comporte un élément « détection » et un élément « identification ». Cela signifie que le décideur est confronté à une incertitude en ce qui concerne le site de l’écrasement, et qu’il doit identifier le site parmi d’autres structures d’aspect similaire sur le terrain. La recherche sur les tâches décisionnelles complexes (détection plus identification) a l’avantage d’être plus en phase avec le monde réel, mais en contrepartie, il y a le coût de la complexité accrue de l’analyse et de la compréhension des données. Étant donné que les tâches décisionnelles complexes font intervenir plusieurs facteurs susceptibles d’influencer la prise de décision, il faut une méthode formelle pour distinguer (clarifier) les effets des divers facteurs, et pour simplifier l’évaluation des résultats de la de prise de décision. Le rapport présente un modèle de tâches décisionnelles complexes (SDT−CD) qui joue précisément ce rôle. Ce modèle a été évalué grâce à l’analyse de plusieurs ensembles de données de démonstration provenant d’une grande variété de domaines, et il a fait la preuve qu’il est capable de dénouer des tâches complexes et de fournir des résultats faciles à interpréter. En plus de mesurer la performance et la partialité du processus de prise de décision, le modèle peut être utilisé pour tester des hypothèses et pour évaluer si la prise de

décision est optimale.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

(U) Decision making; Compound Decision; Detection; Identification; Signal Detection Theory; Modeling; Naturalistic Decision Environment; Search and Rescue; Eyewitness Identification; Tumour Detection

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE DÉFENSE

www.drdc-rddc.gc.ca